Department of Economics

DISCUSSION PAPER SERIES IN ECONOMICS

DP No. 08

Voting Expressively

January 2019

Abhinash Borah

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region
P.O.Rai,Sonepat Haryana-131029

Voting Expressively

Abhinash Borah*

January 28, 2019

Abstract

We address a common criticism directed towards models of expressive voting that they are ad hoc in nature. To that end, we propose a foundation for expressive behavior that is based on a novel theory of social preferences under risk. Under our proposal, expressive considerations in behavior arise from the particular way in which risky social prospects are assessed by decision makers who want to interpret their choices as moral. To illustrate the scope of our framework, we use it to address some key questions in the literature on expressive voting: why, for expressive considerations, might voters vote against their self-interest in large elections and why might such elections exhibit a moral bias (Feddersen et al. 2009). Specifically, we consider an electoral set-up with two alternatives and explain why, when the size of the electorate is large, voters may want to vote for the alternative they deem morally superior even if this alternative happens to be strictly less preferred, in an all-inclusive sense, than the other.

Keywords: expressive voting, morals, social preferences, decisions under risk, voting against self-interest, moral bias of large elections

JEL classification: D01, D03, D81, D72, A13

*Department of Economics, Ashoka University. Email: abhinash.borah@ashoka.edu.in. Address: National Capital Region, P.O. Rai, Sonepat Haryana-131029, India. I am deeply indebted to Andrew Postlewaite and Alvaro Sandroni for several illuminating conversations on the subject of expressive voting.

1 Introduction

In many contexts, especially ones with a social dimension, choices are very often driven not just by an *instrumental* motivation pertaining to outcomes but also by an *expressive* concern that is more directly and intrinsically connected to the very nature of the choices under consideration. This connection draws on how such choices may express features like a decision maker's values, identity and self-image. Consider, for instance, voting behavior in large elections, which is our focus in this paper and is taken as a leading example in the literature of such expressive concerns. In such elections, the probability that a voter's vote is pivotal is fairly insignificant. Therefore, in the presence of even minimal costs associated with voting, voter participation in such elections should be negligible, if such participation is driven purely by an instrumental concern to influence the outcome of the election. As such, non-trivial voter participation, as is seen in real-world elections, suggests that expressive considerations—e.g., a desire to express one's values, morals or ideology, beyond influencing electoral outcomes—may play a significant part in shaping voters' electoral choices.[1]

At this level of generality, the logic of expressive choices seems quite compelling. However, a common criticism against it is that it is ad hoc. Most standard approaches to modelling expressive choices do so by simply adding an *exogenous* non-instrumental utility term to the decision maker's utility function. For instance, in the context of voting such a non-instrumental utility term is very often added to capture the idea that voters find the very act of voting or, in other cases, the act of voting for certain alternatives expressively compelling. However, no account is generally provided for the micro-foundations of this non-instrumental utility term. As such, an implicit assumption that this approach maintains is that the expressive concern that is modelled is somehow particular to the specific problem like voting being studied and is not a deep feature of preferences that applies more generally to other problems of a similar nature.

In this paper, to address this criticism, we propose a foundation for expressive choices that is based on a novel theory of social preferences under risk. Our theory

[1]The theme of expressive voting dates at least as far back as Buchanan (1954) and has been developed further by Downs (1957), Riker and Ordeshook (1968), Tullock (1971), Fiorina (1976) and Brennan and Buchanan (1984), among others. For a comprehensive account on the subject, see Brennan and Lomasky (1993) and Schuessler (2000).

considers decision makers who want to interpret their choices as moral and it models expressive assessments as an integral component of how moral assessments are made. Because of this, in our theorizing, an expressive assessment is not something that remains specific to a particular problem like voting, but is rather a property of a certain class of social preferences under risk, in general. Indeed, we show how in such problems, expressive concerns may *endogenously* emerge as the driving force behind certain choices from the very way in which risky social prospects are assessed by the kind of decision makers we model.

The key idea that our decision model highlights is that when it comes to morality, end-outcomes may not be the only criterion by which moral assessments are made. It is possible that an act may be viewed as a moral one even if it fails to realize a desirable moral end.[2] Specifically, our model considers decision makers who want to perceive their choices in social environments as moral or prosocial. Clearly, if the end-outcome that realizes as a consequence of such a decision maker's choice is favorable from a social point of view, then it allows her to think of her choice as moral. However, beyond this *instrumental view of morality*, she may also concern herself with whether her choice per se, distinct from the particular outcome that realizes from it, can be *interpreted* as moral. Under this *expressive view of morality*, she may want to reason about how favorable, from a social point of view, was her choice, in general, in terms of the overall possibilities under it. As such, if she can reason, for instance, that overall, her choice did allow for a socially more favorable outcome to realize than what actually did realize, then she may, ex post, expressively interpret her choice as more moral than what the resulting outcome warrants. Conversely, if the DM perceives that, overall, her choice actually warranted a socially less favorable outcome to realize than what actually did, then she may be forced to interpret her choice as less moral than what the resulting outcome warrants. That is, we propose that the DM may assess the morality of her choices both from an instrumental as well as an expressive point of view.

Specifically, social situations of risk make both the instrumental and expressive dimension salient. This is because, since chance intermediates between what a decision maker chooses and the outcome that realizes, it allows her to weigh both her overall choice and the outcome that realizes from it as distinct considerations

[2]For instance, think of an individual who risks his own life to save the life of a Muslim man, threatened by Islamophobic violence. Presumably, such an act will be considered a moral one even if it fails to realize its desired moral end of saving the man.

on her moral scale. The key analytical detail underlying the decision model is its take on how instrumental moral assessments are different from expressive ones. Whereas the former is tied to the logic of Bayesian rationality, the latter is not. We formalize this departure of expressive moral reasoning from Bayesian rationality by drawing on the concept of a *probability weighting function* that transforms objective probabilities into subjective decision weights. Specifically, expressive moral assessments may involve a decision maker over-weighting or under-weighting the true probabilities of events. For instance, she may over-weight small chances of a morally favorable outcome occurring, since even a small chance of such an outcome occurring may feel qualitatively much better than there being no chance of it occurring. At the same time, there may be an appeal tied to a morally favorable outcome occurring for sure and even small departures from certainty may be unattractive to her, captured by the under-weighting of such large chances. This phenomenon of over-weighting small probabilities and under-weighting large ones, which is referred to as regressive probability weighting Kahneman and Tversky (1979), plays a crucial part in the way we model expressive moral assessments.

As an illustration of the unified framework that we are proposing, we use our theory of social preferences to relate expressive voting to evidence on prosocial behavior under risk. Then, to show the scope of our framework, we use it to address some key questions in the literature on expressive voting. First, we show why, for expressive considerations, voters might vote against their self-interest. Second, we show why large elections may exhibit a moral bias, i.e., "controlling for the distribution of preferences within the electorate, alternatives that are understood by voters to be morally superior are more likely to win in large elections than in small ones" Feddersen, Gailmard, and Sandroni (2009). Specifically, we consider an electoral set-up with two alternatives and explain, within a rational framework, why, when the size of the electorate is large, voters may want to vote for the alternative they deem morally superior even if this alternative happens to be *strictly* less preferred, in an *all-inclusive* sense, than the other. In so doing, our framework also provides an explanation for the phenomenon of inefficient unanimity—a situation where everyone in the electorate unanimously prefers one alternative to another, but at the same time everyone votes for the less preferred alternative Brennan and Lomasky (1984). Indeed, we show that this situation can prevail even when every voter's vote is guaranteed to be pivotal with positive probability (as is the case under the electoral mechanism that we adopt) and not just when, in equilibrium, there is a zero probability of their vote being pivotal, which is what the argument of Brennan and Lomasky (1984) relies on.

Before we get to the formal details of our decision model, we would like to begin by appealing to some experimental evidence with the goal of motivating the unified framework that we are proposing.

2 Some Motivating Experimental Evidence

Consider the following two experimental findings with the goal of identifying the common thread that runs through them.

Voting Against Self-Interest and Moral Bias of Elections: Feddersen, Gailmard, and Sandroni (2009)—FGS hereafter—consider an experimental election setting in which a group has to choose between two alternatives, A and B. The group is composed of A types and B types. Only B types, numbering n, have a chance of influencing the electoral outcome. B types simultaneously and privately choose one of three options: abstain, vote for A, or vote for B, with voting being costly. After these choices are made, the electoral outcome is determined by selecting one B type individual at random. If the selected individual has voted, then her chosen alternative is the electoral outcome, else the outcome is determined by the flip of a fair coin.

The experimental design allows FGS to implement a decision model of an election in which they can directly control the probability of a voter (i.e., a B type) being pivotal in the election. Specifically, in any treatment, this probability (pivot probability, for short) is given by $\frac{1}{n}$. In their experiment, monetary payoffs are such that under alternative B, B types receive a higher payoff than A types whereas under alternative A both types receive moderate, egalitarian payoffs. Specifically, alternative A minimizes inequality in terms of monetary rewards, maximizes the sum of monetary rewards and maximizes the minimum reward. For these reasons, FGS think of A as the *moral* alternative in this election and B as the *selfish* one (for the voters). By varying the treatment, the key questions that they seek to address is about the impact that changing the pivot probability has on (i) a voter's choice at the individual level, in particular, her propensity to vote for the moral alternative A and (ii) the outcome of the election at the group level.

As far as the first question goes, FGS find that a voter's choice may indeed depend on the pivot probability. Specifically, conditional on voting, voters show a systematic tendency to choose the selfish alternative B when the pivot probability is high but to switch to the moral alternative A when this probability falls. What's noteworthy is that such a pattern of choices cannot be explained by instrumental considerations alone. To understand why, consider a voter who votes for alternative B when she is a dictator, i.e., her vote is pivotal with probability one. Such a voter reveals a strict preference for B over A. If instrumental considerations are all that mattered, then, to be consistent with her revealed preference over these alternatives, in any treatment-election in which she does vote, she should do so in favor of B irrespective of how large the electorate is as long as her pivot probability is positive, as is the case in these experiments. Doing so *strictly* increases the probability of her preferred alternative winning. Voting for A, on the other hand, strictly reduces this probability and goes against her revealed preference over these alternatives. This, however, does not necessarily mean that such choices are irrational. As we have discussed above, considerations of an expressive nature that go beyond the sole concern of influencing electoral outcomes may matter to voters and such expressive considerations may motivate choices of this type.

As far as FGS's second question goes, they find that the influence that the pivot probability has on voters' choices at the individual level carry over to the electoral outcome at the group level. The data shows that as the size of the electorate increases and the pivot probability decreases, the probability of the moral alternative A being the electoral outcome increases and vice versa. That is why FGS conclude that their experimental findings are suggestive of the fact that large elections may exhibit a moral bias.

The Probabilistic Dictator Game: Next, consider the probabilistic dictator (PD) game. In such a game, the dictator is endowed with a fixed amount of money. She is not allowed to share the money with the other individual, but she is given the option, if she so chooses, to share *chances* of getting the money with him. In particular, she can assign him any probability of getting the entire amount while retaining the amount herself with complementary probability. For example, if the fixed amount is \$20 and the decision maker (DM) assigns to the other person a probability $\alpha \in [0, 1]$, then the allocation $(0, 20)$ in which the other person gets the 20 dollars (and the DM gets 0) results with probability α and the allocation $(20, 0)$ in which the DM gets the 20 dollars (and the other person gets 0) results with probability $1-\alpha$. Experimental evidence Krawczyk and LeLec (2010); Brock, Lange, and Ozbay (2013) indicate that a significant portion of decision makers do give the other individual a small positive probability of getting the money.

Here too, the DM's choice cannot be explained in terms of instrumental considerations alone. To see this, consider a DM who chooses to give the other person, say, a 10% chance of receiving the entire amount, despite strictly preferring the allocation $(20, 0)$ to the allocation $(0, 20)$.[3] Imagine that the DM starts off with the lottery that results in the allocation $(20, 0)$ with probability 0.9 and the allocation $(0, 20)$ with probability 0.1. Now, consider replacing this lottery with the degenerate lottery in which she gets the money for sure, i.e., the allocation $(20, 0)$ realizes with unit probability. Observe that what we have done under this replacement is move a 10% chance away from her less preferred outcome—the allocation $(0, 20)$—and put it on her more preferred outcome—the allocation $(20, 0)$—while with 90% chance the outcome remains the same. If the DM were motivated by instrumental considerations alone, then she should be made strictly better off by this replacement as it increases the chance of her preferred outcome realizing. However, given that she chooses the first lottery over the second, the DM that we are considering is made strictly worse off.

Further, observe that the voting pattern in the FGS experiment has a striking similarity to the choices made in the PD game. Clearly, many of the voters care about the well being of others, in particular, the A type individuals, who have the maximum riding on the voters' choices. However, given that they care about their own outcomes as well and the difference between the two alternatives on this dimension, they have reason to strictly prefer the selfish alternative B over the moral alternative A, just as in the PD game, most decision makers would strictly prefer the allocation $(20, 0)$ over the allocation $(0, 20)$ even though they may care about the other individual. Therefore, when any such decision maker is a dictator, or more generally, when the probability that her vote is pivotal is high, she finds it optimum to vote for B. However, when the pivot probability is low, voting for A, it would seem, is like contributing a small chance to the moral alternative winning, and as in the PD game where decision makers do not mind giving a small chance to the other individual getting the money, so too appears to

[3]The content of the argument that we make below continues to hold if we consider the less plausible case of a decision maker who strictly prefers the allocation $(0, 20)$ to the allocation $(20, 0)$.

be the case here. What's different, of course, is the fact that the outcome in the electoral setting, unlike in the PD game, is determined based on an aggregation of individual choices. Not withstanding this difference, what we will show is that the same expressive concern for morality, which characterizes risky social assessments under our decision model, can explain not only the choices in the PD game, but also plays the key role in explaining why voters may vote against their self-interest and why large elections may exhibit a moral bias.

The rest of the paper is organized as follows. Section 3 lays out our decision model of social preferences under risk that formalizes the interaction between instrumental and expressive concerns. To illustrate this interaction, we first show how it rationalizes the choices in the PD game. Then, in section 4, we introduce a voting model and show how a similar interaction explains the phenomenon of voting against self-interest and moral bias of large elections. Section 5 extends the basic voting model to establish the robustness of our main finding. It also illustrates why, in relation to the available literature, the behavioral implications of our social preferences model is novel and essential to explaining the type of expressive voting behavior we study here. Section 6 provides some comments on the literature. Finally, Section 7 concludes. Proofs of the Propositions are provided in the Appendix.

3 A Model of Social Preferences under Risk

Let X denote an underlying set of outcomes with typical element x. We think of these outcomes as ones with a social dimension, possibly impacting the well-being of all members of a given society. For example, the set X could be the set of allocations of private goods, different levels of a public good, possible amounts of public debt, a set of electoral outcomes, etc.[4] $\Delta(X)$ will denote the set of simple probability measures (lotteries, for short) on X with typical element p. In the way of notation, for any $p \in \Delta(X)$, $p(x)$ will denote the probability that p assigns to $x \in X$. Further, for any $x \in X$, δ_x will denote the degenerate lottery that gives x with unit probability.

[4] In keeping with terminology from social choice theory, X may be thought of as the set of social states.

Consider a decision maker (DM) living in this society who has preferences $\succcurlyeq$ on $\Delta(X)$.[5] We now formally introduce the representation of the DM's preferences that we are proposing. The representation captures the idea that the DM's assessment of risky social prospects in $\Delta(X)$ is based on an interaction between selfish and moral concerns, with morality assessed both instrumentally as well as expressively. We call this representation the expressively moral (EM) representation.

Definition 3.1. *An EM representation of $\succcurlyeq$ consists of (i) a function $u : X \to \mathbb{R}$, (ii) a function $V : \Delta(X) \to \mathbb{R}$, with the function $v : X \to \mathbb{R}$ defined by $v(x) = V(\delta_x)$, and (iii) a constant $\sigma \in [0, 1]$, such that the function $W : \Delta(X) \to \mathbb{R}$, given by*

$$W(p) = \sum_{x \in X} p(x) \{u(x) + v(x) + \sigma[V(p) - v(x)]\},$$

represents $\succcurlyeq$.

Under the representation, the function u has the interpretation that it captures the DM's ranking of outcomes in X from a selfish point of view. On the other hand, the functions V and v capture, respectively, her ranking of lotteries in $\Delta(X)$ and outcomes in X from a moral perspective. We maintain that this moral ranking of lotteries/outcomes is based on how prosocial she considers them to be. That is, it is based on how favorable she finds these alternatives from the perspective of the other members of society. For instance, consider the allocations $(20, 0)$ and $(0, 20)$ in the PD game. From a selfish perspective, the DM's assessment would presumably be $u(20, 0) > u(0, 20)$. On the other hand, a DM who wants to perceive herself as moral in the sense maintained here, would most likely find the allocation $(0, 20)$ more prosocial than $(20, 0)$ and her moral assessment would then be given by $v(20, 0) < v(0, 20)$. In other words, morality here reflects the ability to view things from the perspective of what's favorable for others.

Now, consider the DM's evaluation of a lottery p. To understand this evaluation, focus on her assessment in the event that the outcome x in the support of p realizes, given by $u(x) + v(x) + \sigma[V(p) - v(x)]$. This assessment can be broken down into an instrumental component, $u(x)+v(x)$, and an expressive component, $V(p) - v(x)$, with $\sigma \in [0, 1]$ the weight put on the latter. The instrumental component is standard. It captures the DM's comprehensive—selfish as well as moral—instrumental assessment of the outcome x. The distinguishing feature

[5] Formally, $\succcurlyeq \subseteq \Delta(X) \times \Delta(X)$.

of our decision model, as discussed earlier, is that moral assessments are not restricted to an instrumental consideration only. It may include an expressive consideration as well. The expressive consideration captures the idea that the DM may care about whether her choice per se, distinct from the particular outcome that realized from it, can be *interpreted* as moral or not. Under this expressive view of morality, even ex post, after uncertainty about the outcome has been resolved, she may still be concerned with the question of how favorable was her overall choice from a social point of view. Specifically, in the event that outcome x realizes, if the DM considers this outcome to be (morally) inferior than her overall moral assessment of the possibilities under p, i.e., $V(p) > v(x)$, then she can reason that, overall, her choice did allow for a socially more favorable outcome to realize than what actually did realize. She may, therefore, expressively interpret her choice, ex post, as more moral than what the resulting outcome warrants. To reflect this, a positive adjustment of $\sigma[V(p) - v(x)]$ is added to the instrumental assessment. On the other hand, if $v(x) > V(p)$, then the DM perceives that, overall, her choice actually warranted a socially less favorable outcome to realize than what actually did realize and she is, therefore, compelled to interpret her choice as less moral than what the resulting outcome warrants. To reflect this, a negative adjustment of $\sigma[V(p) - v(x)]$ is added to the instrumental assessment. The DM's evaluation of the lottery p is then given by integrating, with respect to the probabilities, her assessments under the different events associated with the realization of the different outcomes in the support of p.

As the reader must have noted, our decision model critically depends on the properties of the function V. First, observe that if this function is linear in probabilities, then the assessment of any lottery p reduces to:

$$W(p) = \sum_{x \in X} p(x)(u(x) + v(x)),$$

i.e., an expected utility assessment of p with a Bernoulli utility function that is additive across the DM's selfish and moral instrumental concerns. The substantive content of our decision model, therefore, is seen when V is non-linear in probabilities. It is this non-linearity, which marks a departure from Bayesian rationality, that captures the essence of expressive reasoning. For the purpose of this paper, we restrict attention to the specification of the V function for lotteries in $\Delta(X)$ with two elements in its support. For any such lottery $[x, \mu; x', 1-\mu]$,[6] with $v(x)$

[6] Following standard notation, $[x, \mu; x', 1-\mu]$ denotes the lottery that gives x with probability μ and x' with probability $1-\mu$.

$\geq v(x')$, we assume that the function V takes a bi-separable form:[7]

$$V([x,\mu;x',1-\mu]) = \varphi(\mu)v(x) + (1-\varphi(\mu))v(x'),$$

where $\varphi : [0,1] \to [0,1]$ is a *probability weighting function*, i.e., a strictly increasing function, with $\varphi(0) = 0$ and $\varphi(1) = 1$. The probability weighting function has the interpretation that it transforms objective probabilities into decision weights. In the current context, these decision weights capture the attitude that the DM has from a moral perspective toward social risks. For instance, consider a lottery that results in a socially favorable outcome with a 95% chance and a socially unfavorable outcome with a 5% chance. In this case, from her moral perspective, the DM may not consider a 95% *chance* of the favorable outcome realizing as equivalent to 95% of the *value* of this outcome realizing for sure, as a linear v function would imply (Wakker (2010), pg. 147). At a psychological level, the two assessments may feel very different to her and the probability weighting function is meant to capture this difference. In this paper, we hypothesize that the DM's moral assessments may overweight small chances of a socially favorable outcome realizing, because, to her, such an outcome realizing with some positive chance, even if it happens to be a really small one, may qualitatively feel much better than their being no chance of it realizing. At the same time, her moral assessments may underweight large chances of a socially favorable outcome realizing, because, to her, such an outcome realizing for sure may feel qualitatively much better than their only being a chance of this outcome realizing, even if that chance is a really big one. This phenomenon of over-weighting small probabilities and under-weighting large ones, which is referred to as *regressive probability weighting*, has been extensively documented in the context of decision makers' risk attitudes toward "private lotteries," going at least as far back as the important contribution of Kahneman and Tversky (1979). In this paper, we entertain the possibility that similar effects may be at play when a decision maker assesses social risks from a moral perspective.

It is worth pointing out that the particular form that expressive assessments take in our model may be, following Sandbu (2007), described as *dependent non-instrumental valuation.* That is, valuation that depends on its standing in a par-

[7]As the name suggests, the bi-separable form introduces event-separability in the moral assessment of lotteries in $\Delta(X)$ in a very limited sense, i.e., lotteries that put positive probability on only two outcomes are assessed by the DM in the spirit of generalized expected utility. The bi-separable form is a special case of a rank dependent specification. Ghirardato and Marinacci (2002) consider bi-separable preferences in the context of uncertainty.

ticular relation to end-outcomes, but at the same time is not exclusively reducible to its causal efficacy in bringing about end-outcomes. Sandbu associates such valuation with the *symbolic* aspect of choices or actions, which is similar in spirit to how the DM that we model concerns herself with the question of whether her choice per se, distinct from the particular outcome that realizes from it, can be *interpreted* as moral or not.

A final aspect of the model that is worth emphasizing is that, because of the expressive consideration that it incorporates, it can accommodate violations of stochastic dominance. Formally:

Definition 3.2. *A lottery $p' \in \Delta(X)$ first-order stochastically dominates another lottery $p \in \Delta(X)$ with respect to the DM's preferences if for all $x \in X$, the probability that p' assigns to outcomes that are better than x is at least as great as the corresponding probability under p, and for some $x \in X$ it is strictly greater. DM's preferences satisfy* ***stochastic dominance*** *if whenever p' first-order stochastically dominates p, then $p' \succ p$.*

Stochastic dominance is a natural generalization of the idea that "more (of a good) is better" to environments of risk. It requires a decision maker's assessments to conform to a basic form of consequentialism, namely, that her ranking of outcomes should be independent of the lotteries generating these outcomes. However, that is precisely the kind of event-wise preference separability that the DMs that we model find hard to adhere to. This is because these DMs are motivated by expressive moral considerations and the lottery generating end-outcomes informs them about whether they can interpret their choice per se, distinct from the outcome that realized from it, as moral or not. In essence, the property of stochastic dominance conveys the message that only instrumental considerations should matter to DMs. That is why, violations of stochastic dominance are quite common for decision makers with social preferences. For example, the preferences of the experimental subjects in the probabilistic dictator game discussed above violates stochastic dominance. We now go back to this evidence to further illustrate the working of our decision model.

3.1 Choices in the Probabilistic Dictator Game

We now highlight how a tradeoff between instrumental and expressive considerations, which our decision model embeds, explains the choices that we described in the probabilistic dictator game. So, consider a DM who strictly prefers the allocation $(20,0)$ to the allocation $(0,20)$ on selfish grounds but strictly prefers the latter allocation to the former on moral grounds, i.e.,

$$\overline{u} := u(20,0) > u(0,20) =: \underline{u} \text{ and } \overline{v} := v(0,20) > v(20,0) =: \underline{v}$$

Overall, the DM strictly prefers the allocation $(20,0)$ to $(0,20)$, i.e.,

$$\overline{u} + \underline{v} = u(20,0) + v(20,0) > u(0,20) + v(0,20) = \underline{u} + \overline{v}.$$

Now, consider the DM's problem of deciding what probability $\lambda \in [0,1]$ she wants to assign to the other individual of getting the 20 dollars in this game. Any choice of λ generates a lottery, $p(\lambda) = [(0,20),\ \lambda;\ (20,0),\ 1-\lambda]$, under which the allocation $(0,20)$ is realized with probability λ and the allocation $(20,0)$ is realized with probability $1-\lambda$. If the DM's preferences have an EM representation, then her evaluation of any such lottery $p(\lambda)$ is given by:

$$W(p(\lambda)) = \lambda\{\underline{u} + \overline{v} + \sigma[V(p(\lambda)) - \overline{v}]\} + (1-\lambda)\{\overline{u} + \underline{v} + \sigma[V(p(\lambda)) - \underline{v}]\}$$

Further, if V takes a bi-separable form, i.e.,

$$V([(0,20),\lambda;(20,0),1-\lambda]) = \varphi(\lambda)\overline{v} + (1-\varphi(\lambda))\underline{v},$$

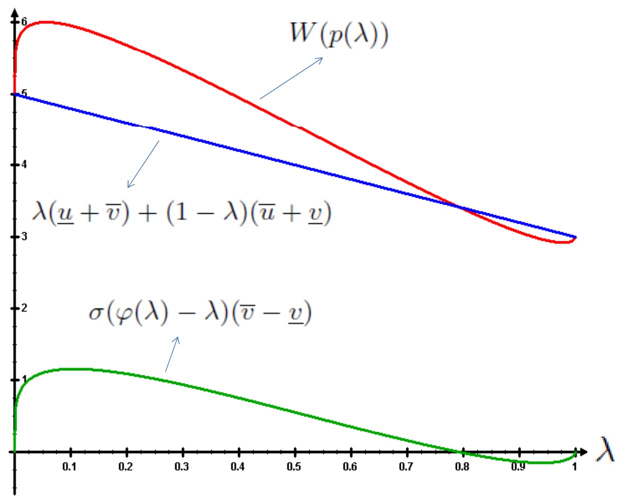
where $\varphi : [0,1] \to [0,1]$ is a probability weighting function, then we have:

$$W(p(\lambda)) = \lambda(\underline{u} + \overline{v}) + (1-\lambda)(\overline{u} + \underline{v}) + \sigma(\varphi(\lambda) - \lambda)(\overline{v} - \underline{v}).$$

Assume that the probability weighting function is continuous, differentiable and takes a regressive form. In that case, the situation looks like the one depicted in figure 1.[8] Note that the choice of λ influences both the instrumental assessment, $\lambda(\underline{u}+\overline{v}) + (1-\lambda)(\overline{u}+\underline{v})$, as well as the expressive assessment, $\sigma(\varphi(\lambda)-\lambda)(\overline{v}-\underline{v})$. Under our assumption that the DM strictly prefers the allocation $(20,0)$ to the

[8]The specification of the probability weighting function we use for the figure is given by $\varphi(\lambda) = \exp(-\beta(-\ln\lambda)^{\alpha})$, where we assume that $\alpha = 0.65$ and $\beta = 0.6$. This particular class of probability weighting functions has been axiomatized by Prelec (1998). Further, we assume that $\sigma = \frac{3}{4}$

Figure 1: Choices in the Probabilistic Dictator Game.



allocation $(0, 20)$, i.e., $\overline{u} + \underline{v} > \underline{u} + \overline{v}$, the instrumental payoff decreases linearly in λ. On the other hand, since $\overline{v} > \underline{v}$, the expressive payoff first increases, attains a maximum, and then starts decreasing with λ (before increasing again). Accordingly, as figure 1 illustrates, for values of λ less than the maxima of the expressive payoff a tradeoff emerges: by her instrumental consideration, as λ increases, she becomes worse off, but by her expressive consideration, she is better off. The change in the overall payoff is determined by the interaction of these two opposing influences. In particular, if φ is sufficiently steep in the neighborhood of 0, i.e.,

$$\lim_{\lambda \to 0^+} \varphi'(\lambda) > 1 + \frac{1}{\sigma}\left(\frac{\overline{u} - \underline{u}}{\overline{v} - \underline{v}} - 1\right),$$

then for small values of λ, on increasing λ slightly, the improvement in the expressive payoff outweighs the drop-off in the instrumental payoff. That is why the payoff $W(p(\lambda))$ from the lottery $p(\lambda) = [(0, 20),\lambda;\ (20, 0),1-\lambda]$ may be increasing in λ and such a DM may choose to give the other individual a positive probability of getting the money, even though she strictly prefers the allocation $(20, 0)$ to the allocation $(0, 20)$.

Now that we have laid out the basics of our decision model, in particular, the interaction between instrumental and expressive considerations that it embeds, we use it to address the issue of expressive voting and, in the process, illustrate the unified framework that we are proposing. Specifically, we show why in large elections voters might vote against their self-interest and why such elections may exhibit a moral bias.

4 Voting Model

Consider a polity consisting of n voters that has to decide between two alternatives, A and B. To fix ideas, think of A and B as two fiscal policies. Alternative A is the policy of running a balanced budget in the current time and alternative B is the policy of running budget deficits at the current time and accumulating debt, which will have to be repaid by future generations. Assume that the deficits under policy B are incurred to provide subsidies to all the voters whereas balanced budgets under A are made possible by avoiding these. So, from the perspective of a typical voter's selfish consideration, alternative B may be preferable to alternative A. But, at the same time, such a voter may find the act of borrowing from unborn future generations to pay for these subsidies highly irresponsible. So, from her moral perspective alternative A may be preferable.

We assume that voting is costless so that everyone in the electorate votes in the election. Finally, the outcome of the election is determined by a mechanism that mimics the one used by FGS. First, all voters simultaneously and privately choose whether to vote for alternative A or B. After all voters have reported their choice, one voter is selected at random, and her chosen alternative becomes the electoral outcome. Accordingly, under this mechanism, the probability that any voter's vote is pivotal is $\frac{1}{n}$. The last two modeling assumptions are made primarily with the goal of isolating in the clearest possible terms the critical decision-theoretic tradeoff between instrumental and expressive concerns that is our focus here. Our arguments can be replicated in a more complicated environment that involves features like costly voting, private information and plurality rule as the mechanism determining the outcome of the election (see Section 5).

The size of the electorate, n, is the crucial variable of interest in our model as our goal is to identify how the voters' choices and the electoral outcome varies with

n. In order to analyze this, we need to ensure that the distribution of preferences within the electorate stays fixed as we vary n. The simplest way to achieve this is to assume that all individuals in the electorate have identical preferences. In particular, we assume that each of them *strictly* prefers alternative B to alternative A. If that is so and, in addition, each of their preferences satisfies stochastic dominance, then, the following conclusion follows immediately.

Proposition 4.1. *If each voter strictly prefers alternative B to alternative A and each of their preferences satisfies stochastic dominance, then for any n, there exists a unique Nash equilibrium (in dominant strategies) in which every voter votes for alternative B. That is, voters' choices and the electoral outcome are independent of the size of the electorate.*

As mentioned earlier, the property of stochastic dominance emphasizes the idea that only instrumental considerations should matter to a DM. Therefore, in a two alternative election, the voting behavior of a DM who adheres to this property is driven exclusively by the goal of maximizing the probability of her preferred alternative winning. We now build on our decision model to show why, when, in addition, expressive concerns matter to a voter, she may, in a completely rational way, end up voting for her less preferred alternative A in a large election.

Observe that since the preferences of all voters are assumed to be identical, we can specify the preferences of the voters in terms of a single *representative voter* (RV). To model expressive concerns, we assume that the RV's preferences can be represented by an EM representation. Specifically, her assessment of any lottery $p = [A, \mu; B, 1 - \mu]$ over electoral outcomes is given by:

$$W(p) = \mu\{u(A)+v(A)+\sigma[V(p)-v(A)]\}+(1-\mu)\{u(B)+v(B)+\sigma[V(p)-v(B)]\}$$

We assume that the RV considers alternative B to be strictly preferable from the perspective of her selfish concern, i.e.,

$$u_H := u(B) > u(A) =: u_L.$$

On the other hand, she considers alternative A to be strictly preferable from the perspective of her moral concern, i.e.,

$$v_H := v(A) > v(B) =: v_L.$$

Overall, as mentioned earlier, she strictly prefers alternative B to alternative A, i.e.,

$$u(B) + v(B) = u_H + v_L > u_L + v_H = u(A) + v(A).$$

In the way of notation, let $\nu = \frac{u_H - u_L}{v_H - v_L}$. Clearly, the above assumption implies that $\nu > 1$. Further, the function V takes a bi-separable form. This means that there exists a probability weighting function $\varphi : [0,1] \to [0,1]$ such that the RV's moral assessment of a lottery $[A, \mu; B, 1-\mu]$ over electoral outcomes is given by:

$$V([A, \mu; B, 1-\mu]) = \varphi(\mu)v_H + (1 - \varphi(\mu))v_L.$$

Finally, we assume that:

- $[A1]$ The probability weighting function φ is differentiable on $[0,1]$ and there exists $\hat{\mu} > 0$ such that for all $\mu \in [0, \hat{\mu}] \cup [1-\hat{\mu}, 1]$, $\varphi'(\mu) > 1 + \frac{\nu-1}{\sigma}$.[9]

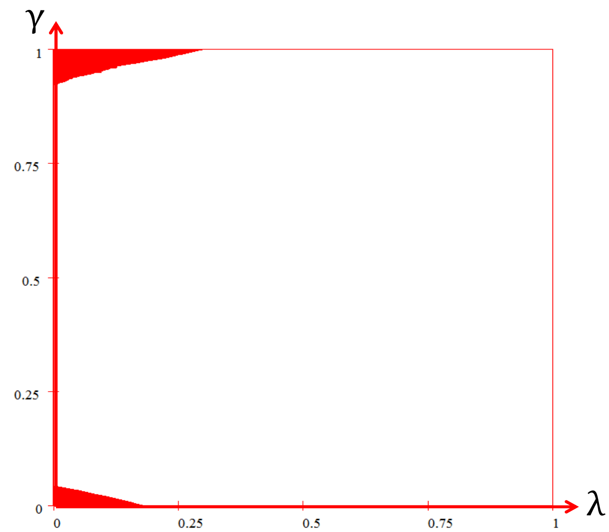
Assumption $[A1]$ plays a critical role in our analysis. Given that $\nu > 1$, it implies that for all $\mu \in [0, \hat{\mu}] \cup [1-\hat{\mu}, 1]$, $\varphi'(\mu) > 1$. It follows that there exists a neighborhood of 0 in which $\varphi(\mu) > \mu$, and a neighborhood of 1 in which $\varphi(\mu) < \mu$. In particular, in the context of the voting problem, this means that the RV's moral assessment over-weights small probabilities and under-weights large ones of her morally preferred outcome A being realized.

We can now state our central result. It establishes that even though the RV strictly prefers alternative B to alternative A, it is perfectly consistent with rational, preference-maximizing behavior for her to vote for alternative A in large elections owing to expressive considerations. Further, the hypothesis of a moral bias in large elections is simply an equilibrium consequence of such voting behavior.

Proposition 4.2. *(Voting Against Self-Interest and Moral Bias) Suppose assumption* $[A1]$ *holds. Then:*

[9] Differentiability of φ at 0 and 1 is defined as $\varphi'(0) := \lim_{\mu \to 0^+} \varphi'(\mu)$ and $\varphi'(1) := \lim_{\mu \to 1^-} \varphi'(\mu)$, respectively. The assumption is stronger than what we need to prove our main result below. We choose to stick to this assumption as most probability weighting functions that are considered in the literature take a differentiable form. For instance, the probability weighting function suggested by Prelec (1998) that was mentioned in an earlier footnote satisfies this condition provided the weight σ is positive, which we always assume to be the case.

Figure 2: Comparing Payoffs from Voting for Alternatives A and B



1. *There exists a positive integer $\overline{n}$ such that whenever $n \geq \overline{n}$ everyone voting for alternative A is the unique symmetric pure strategy Nash equilibrium.*

2. *There exists a positive integer $\underline{n}$ such that whenever $n \leq \underline{n}$, everyone voting for alternative B is the unique symmetric pure strategy Nash equilibrium.*

The proof is available in the Appendix. Here, we briefly go over the reasoning that drives the result. Consider figure 2, which has been constructed by taking specific values of u_H, u_L, v_H, v_L and a functional form for the probability weighting function that are consistent with assumption $[A1]$. In the figure, the horizontal axis plots the probability, denoted by λ, that the RV's vote is pivotal, whereas the vertical axis plots the probability, denoted by γ, that alternative A is the electoral outcome when her vote is not pivotal. Of course, γ is an endogenous object.[10] On the other hand, under our electoral mechanism, $\lambda = 1/n$, where n is the size of the electorate. The figure characterizes the RV's decision problem for different values of λ and γ. In particular, the two shaded regions capture the interesting implication of our model—for any (λ, γ) pair in these regions, the RV's

[10]Under the voting mechanism, for any choice of strategy on the part of the other voters, the probability γ is well defined.

best response is to vote for alternative A.[11]

Consider, first, the lower southwest region where both λ and γ are small. In this scenario, alternative B is the likely electoral outcome. The RV could add to this likelihood by voting for this alternative, and if she did, she would be made strictly better off on account of her instrumental concern. But, given that the probability λ that her vote is pivotal is small, the increase in payoff on this account is small as well. Thus, her vote is relatively insignificant when viewed from an instrumental perspective. On the other hand, on account of her expressive concern, if she votes for alternative A and increases the chance of this alternative being the electoral outcome, then, on this account, she is made strictly better off. What makes the increase in payoff from this significant (relative to the increase in her instrumental payoff if she votes for alternative B) is the fact that her moral assessment over-weights small chances of her preferred moral outcome, A, being realized. As such, this boost in her expressive payoff from voting for A does not pale into insignificance with a reduction in the pivot probability as fast as the instrumental payoff does. So to sum up, voting for B is almost identical to voting for A from the perspective of her instrumental concern. On the other hand, voting for A is comparatively much better than voting for B from the perspective of her expressive concern. Accordingly, under this scenario, she votes for alternative A.

Next, consider the upper northwest region. In this scenario, given that γ is large, alternative A is the likely electoral outcome. Further, since λ is small, this is true irrespective of which way the RV votes. Therefore, once again, her vote has a minimal impact in terms of influencing her payoff on account of her instrumental concern. On the other hand, voting for A is relatively much better than voting for B on account of her expressive concern. This is true since by voting for A, she can push the probability that this alternative will be the electoral outcome even closer to one. Given that her moral assessment under-weights large chances of her preferred moral outcome, A, being realized, doing so has a favorable impact on her payoff. Once again, this expressive payoff does not decrease with the pivot probability as fast as the instrumental payoff does. Accordingly, under this scenario, too, she votes for alternative A.

Given the structure of payoff differences, it should be obvious that when every-

[11] On the other hand, for any (λ, γ) pair outside these regions, her best response is to vote for alternative B.

one else is voting for alternative A (i.e., $\gamma = 1$), for small pivot probabilities, or equivalently, for large electorates, the RV's best response is to vote for alternative A as well. Hence, everyone voting for A is a Nash equilibrium. This equilibrium illustrates the phenomenon of inefficient unanimity, which refers to a situation where everyone in the electorate unanimously prefers one alternative to another, but at the same time everyone votes for the less preferred alternative Brennan and Lomasky (1984). It is worth noting that, in our analysis, this situation prevails even when every voter's vote is guaranteed to be pivotal with positive probability (as is the case under the electoral mechanism that we adopt). This is different from the analysis in Brennan and Lomasky (1984), which relies on the fact that, in equilibrium, each voter's vote is pivotal with zero probability.

5 Further Remarks on the Model

In this section we first show that the expressive logic that our decision model implies for voting behavior is retained when we extend the voting model developed above to allow for more realistic features like costly voting, private information and using plurality rule as the voting mechanism. This establishes that the result we derived above is not a consequence of abstracting away from these real world features. Second, we answer the question as to why we needed to propose a new model of social preferences under risk to capture the type of expressive voting behavior we highlight here. The reader is well entitled to wonder why we did not take a well established model of social preferences under risk to convey the main message of the paper. To answer this question, we consider the leading model of social preferences under risk in the literature, the expected inequality aversion (EIA) model, that has been proposed by Saito (2013). The EIA model extends the inequality aversion logic of Fehr and Schmidt (1999) to environments of risk where decision makers may care about not just inequality of outcomes, but also about inequality of opportunities. We show that this model and the inequality aversion rationale embedded in it cannot accommodate the type of expressive voting behavior that we are considering here. As such, the model of social preferences under risk we are proposing has novel behavioral implications that are essential to explaining the type of expressive voting behavior we are capturing here. We also establish that this observation goes well beyond the voting problem under consideration and the decision model we are proposing provides the language to incorporate certain expressive motivations underlying socially motivated behavior that the existing literature may not have fully focussed on. Because of this, as we will show, it can capture certain classes of experimental evidence that existing models like the EIA cannot. As a final task in the section, we go back to the FGS model and highlight what the implication of adding probability weighting to that set-up is.

5.1 Costly Voting

We introduce two changes to the model of the last section. First, we assume that voting is costly and individuals may have to incur a cost if they choose to vote. Given this, an individual may find it in her best interest to abstain from voting in order to avoid this cost. In other words, an individual has three alternatives now: vote for A, vote for B or abstain, with the first two requiring her to incur a cost whereas the third does not. Second, the outcome of the election is determined by plurality rule. That is, the winner of the election is the alternative that receives the greater number of votes. In case of a tie, we assume that alternative B is the electoral outcome.

Cost of voting may vary across individuals and we assume that this cost is an individual's private information. In the language of Bayesian games, an individual's cost is her type. We assume that individuals' costs are independently drawn from a commonly known distribution with support $[0, \overline{c}] \subseteq \mathbb{R}$, $\overline{c} > 0$, whose cumulative distribution function is denoted by F. We assume that F has a density f that is positive on $[0, \overline{c}]$ and $F(0) > 0$. The last assumption is made to simplify the analysis. It is innocuous in the sense that $F(0)$ just needs to be positive and can be arbitrarily small.

All voters are of the expressively moral type and identical, except for their cost of voting. We assume that the cost of voting impacts an individual's preferences through her selfish perspective. Specifically, if the electoral outcome is $x \in \{A, B\}$ and her cost of voting is c, then by her selfish perspective, her utility is $u(x) - c$, if she votes; and $u(x)$, if she abstains. That is, these two considerations enter her selfish assessment in a separable manner. This means that if a voter has to evaluate an uncertain scenario reflected by a lottery p over the electoral outcomes, A and B, then this evaluation is $W(p) - c$, if she votes, and $W(p)$, if she abstains, where $W(p)$ refers to her EM assessment of this lottery.[12] We continue to assume that all individuals strictly prefer alternative B to A, i.e., $u(B)+v(B) = u_H+v_L > u_L + v_H = u(A) + v(A)$, and that assumption $[A1]$ holds.

The strategic considerations of this electoral set-up can be captured as a Bayesian game. Formally, an individual's strategy is a mapping, $s : [0, \bar{c}] \to \{vote\ A,\ vote\ B,\ abstain\}$. The following result establishes that the message of our earlier result can be replicated in this setting using the language of a Bayes-Nash equilibrium.

Proposition 5.1. *There exists $\overline{n} \in \mathbb{N}$, such that for all $n \geq \overline{n}$, there exists $c_n^* \in (0, \bar{c})$ such that all voters following the strategy*

$$s(c) = \begin{cases} vote\ A, & if\ c \leq c_n^* \\ abstain, & if\ c > c_n^* \end{cases}$$

is a symmetric Bayes-Nash equilibrium of this n person society.

In other words, for a society of size $n > \overline{n}$, in any such Bayes-Nash equilibrium, all voters with a cost lower than c_n^* vote and all those who vote choose their less preferred alternative A. Accordingly, the probability that A is the electoral outcome, $1 - (1 - F(c_n^*))^n$, approaches 1 in large societies.

5.2 Inequality Aversion and Expressive Voting

The EIA Model To lay out the EIA model, consider, like above, a society with n individuals. $I = \{1, \ldots, n\}$ denotes the set of individuals. A vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ will be referred to as an allocation of payoffs. For any individual, say i, who cares about inequality, Fehr and Schmidt (1999) proposed the following utility function over such allocations:

$$w^{FS}(x) = x_i - \sum_{j \in I \setminus \{i\}} [\alpha \max\{x_j - x_i, 0\} + \beta \max\{x_i - x_j, 0\}]$$

The term $\alpha \max\{x_j - x_i, 0\}$ captures i's disutility from disadvantageous inequality with j; whereas $\beta \max\{x_i - x_j, 0\}$ captures her disutility from advantageous

[12] That is, if $p = [A, \mu; B, 1 - \mu]$ is the lottery under consideration, then:
$W(p) = \mu[u(A) + v(A) + \sigma(V(p) - v(A))] + (1 - \mu)[u(B) + v(B) + \sigma(V(p) - v(B))]$

inequality with j.[13]

Now consider a lottery p over payoff allocations like x above; say, p yields the allocation $x^s = (x_1^s, \dots, x_n^s)$ with probability λ_s, $s \in \{1, \dots, m\}$, $\sum_{s=1}^m \lambda_s = 1$. The key observation underlying the EIA model is that an inequality averse individual may not consider an expected utility evaluation of this lottery w.r.t. the Fehr-Schmidt utility function, $E_p(w^{FS}(x)) := \sum_{s=1}^m \lambda_s w^{FS}(x^s)$, an appropriate assessment of it. This is because along with her concern for inequality of ex post outcomes, she may also care about inequality of ex ante opportunities. However, an expected utility evaluation can accommodate only the former concern and not the latter. To address this issue, the EIA model proposes that we should consider the vector of expected payoffs under a lottery like p, $E_p(x) := \sum_{s=1}^m \lambda_s x^s = (\sum_{s=1}^m \lambda_s x_1^s, \dots, \sum_{s=1}^m \lambda_s x_n^s)$, as indicative of the ex ante opportunities of the different individuals under this lottery. A DM's concern for inequality on this dimension can be assessed using the Fehr-Schmidt utility function and this assessment can, then, be combined with the expected utility evaluation to arrive at her overall assessment of such a lottery. That is, a DM's assessment of a lottery like p, under this model, is given by a weighted average of her concerns for ex ante opportunities and ex post outcomes:

$$W^{EIA}(p) = \rho w^{FS}(E_p(x)) + (1 - \rho) E_p(w^{FS}(x)), \text{ with } \rho \in (0, 1)$$

We now show that the EIA model may not be able to capture the type of expressive voting behavior that we have highlighted in this paper. We demonstrate this by considering a set-up of the type in the FGS experimental election. This set-up is suitable for the EIA model as the electoral alternatives in their experiment are in the form of payoff allocations. So, assume that the population comprises of B-types (voters), numbering n, and A-types (non-voters), numbering n'. In the subsequent analysis, we keep n' fixed, while varying n.[14] Each B-type has to vote for either alternative A or alternative B. To highlight our main point in a transparent manner, we go back to assuming that there are no costs associated with voting and, hence, abstention is not an option. In keeping with the FGS experiment and rest of the paper, we consider alternative B to be more favorable

[13] Typically, α is taken to be greater than β to capture the idea that individuals are more sensitive to disadvantageous inequality. Of course, the parameters α and β can also be allowed to vary across individuals $j \in I \setminus \{i\}$.

[14] Think of A-types as a fixed group of refugees and decisions that pertain to their wellbeing are being taken in electorates of different sizes.

towards voters and A towards non-voters. The outcome of the election is determined by the same electoral mechanism as in Section 4. As such, the probability that any voter's vote is pivotal is $\frac{1}{n}$.

Alternative A gives a payoff of x_A to a voter and y_A to a non-voter. Alternative B gives a payoff of x_B to a voter and y_B to a non-voter. Accordingly, the payoff allocation generated by alternative A can be compactly written as (x_A, y_A) and that by B as (x_B, y_B). Observe that, which ever electoral outcome prevails, from a representative voter's perspective, there is no inequality with any of the other voters. Inequality exists only with non-voters and the magnitude of this inequality is the same w.r.t any such non-voter. We assume that $x_B > x_A \geq y_A > y_B$. This assumption is made because it is consistent with the payoffs in the FGS experiment. We can establish a similar conclusion as the one that we derive here under the assumption that $x_B > y_A > x_A > y_B$, although the derivations for that case are more involved. We assume that all voters are of the EIA type. By this, we mean that any such voter evaluates a lottery like $[(x_A, y_A), \pi; (x_B, y_B), 1 - \pi]$ over electoral outcomes using the function W^{EIA}. We assume that all voters' preferences are identical, meaning that the three parameters that characterize the EIA model, (α, β, ρ) are the same for all voters. Finally, in keeping with the experimental finding and the basic hypothesis of Section 4, we assume that all voters prefer alternative B to alternative A. That is, $w^{FS}(x_B, y_B) > w^{FS}(x_A, y_A)$.

The following result establishes that the phenomenon of voting against self-interest and moral bias of large elections involving a vote switch to alternative A for large electorates is not consistent with the EIA model.

Proposition 5.2. *An EIA voter's choice is independent of n and she always votes for alternative B.*

As such, from a behavioral perspective, our decision model does add something novel when it comes to studying such expressive voting behavior.

5.3 Behavioral Underpinnings

The novel behavioral implications that our decision model provides is not limited to the voting problem. We now highlight this claim and in the process throw further light on some of the emotions underlying socially motivated behavior that can be accommodated within it that are fairly distinct from ones modeled in the literature. The leading paradigm within which socially motivated behavior has been studied in recent years in Economics is the inequality aversion framework proposed by Fehr and Schmidt (1999). The EIA model as we showed above is nestled within this paradigm. The inequality aversion paradigm is built on the two emotions of guilt (towards advantageous inequality) and spite (towards disadvantageous inequality). Whereas these, no doubt, are powerful emotions that drive socially motivated behavior, they need not be the sole drivers of such behavior. The EM model speaks to this. It highlights expressive motivations underlying moral evaluations under which, ex post, the DM cares not just about the moral appeal of the outcome that actually realizes but also about how she may interpret or reason about the morality of her choice per se. In so doing, it opens up the possibility of thinking about socially motivated behavior through the prism of concerns like self-image. Further, what it highlights using the construct of probability weighting is that the DM may have a fair amount of leeway when it comes to interpreting the expressive morality of her choices in a favorable manner.

As a way of illustrating the above observations, let us begin by considering the following example.[15] Consider a decision maker who has to choose between the following two ways of allocating \$20 between herself and another individual. She can either choose to allocate the money to the other person for sure or she may choose to reduce his chance of getting the money for sure by a mere 1% which she can assign to herself receiving the amount. That is, her choice is between the sure allocation $(0, 20)$ and the lottery $[(0, 20), 0.99; (20, 0), 0.01]$. How will she choose in this situation? The logic of inequality aversion, specifically, spite towards disadvantageous inequality, as embodied in the EIA model has a very clear implication in this choice problem. The EIA model predicts that the DM will choose to reduce the other individual's chance of receiving the money by 1% and choose $[(0, 20), 0.99; (20, 0), 0.01]$ over $(0, 20)$. However, we can easily think of DMs who may be willing to forego this small 1% chance of getting the money and let the other individual have the amount for sure. Such a DM may reason that ex post, the other individual will get the money in either case—in one case for sure and in the other almost surely. Therefore, by choosing to reduce his ex ante chance

[15]At this point, this example is in the nature of a thought experiment as we do not know of any experimental evidence that has directly tested it. We hope to test this experimentally in the future.

by 1%, all she manages to do is dilute her ability to *interpret* her choice ex post as one borne out of, say, pure sacrifice, which is clearly an emotion that animates socially motivated behavior. This is precisely the kind of expressive motivation that our model can capture that is very distinct from the inequality aversion paradigm. Figure 1 clearly shows that for the particular parameterizations of the EM model used there, the choice of the sure allocation $(0, 20)$ is precisely what the model predicts as the DM's assessment of this allocation is higher than her assessment of $[(0, 20), 0.99; (20, 0), 0.01]$.[16]

We next turn to some experimental evidence to further highlight the behavioral underpinnings of our model. Brock, Lange, and Ozbay (2013) considered a series of dictator games in which the dictator (the DM) was given the task of allocating 100 tokens between herself and a second player (the recipient). Tokens translated to monetary payments with the exact nature of this translation varying from one task to the other. One of the tasks ($T1$) replicated the ordinary dictator game under certainty. In it, if the DM gave $\theta \geq 0$ tokens to the other individual, then the resulting allocation was $(100-\theta, \theta)$. In another task ($T3$), the tokens that the DM gave translated to a lottery for the other individual. More precisely, if the DM gave θ tokens to the other individual, where θ could be no greater than 50, then her payoff was $100 - \theta$ like in $T1$. On the other hand, the other individual faced the lottery $[50, \frac{\theta}{50}; 0, 1 - \frac{\theta}{50}]$, $\theta \in [0, 50]$. In both these tasks a significant proportion of decision makers gave a positive number of tokens to the recipient. The question that interests us here is about the ranking of these two tasks in terms of the number of tokens given by non-selfish dictators.[17]

To appreciate the novelty that the EM model offers, it is instructive to work out what the answer of the EIA model to this question is. Observe that for a DM whose preferences have an EIA representation, her assessment of the allocation $(100 - \theta, \theta)$ is given by:

$$w^{FS}(100 - \theta, \theta) = 100 - \theta - \alpha \max\{2\theta - 100, 0\} - \beta \max\{100 - 2\theta, 0\}$$

It is straightforward to verify that the optimal θ chosen can never be greater than

[16] Of course, it is an empirical question as to what proportion of individuals choose according to the prediction of the EIA model and what proportion according to our's. At this point, all we want to emphasize is the very distinct predictions made by the two models in this choice problem which, in turn, correspond to very distinct psychological motivations.

[17] Brock, Lange, and Ozbay (2013) consider a dictator to be non-selfish if she gave a positive number of tokens in at least one of their tasks.

50, so that $w^{FS}(100-\theta,\theta) = 100-\theta-\beta(100-2\theta)$ and the optimal allocation rule is specified by:

$$\theta^* = \begin{cases} 0, & \beta < 0.5 \\ \in [0,50], & \beta = 0.5 \\ 50, & \beta > 0.5 \end{cases}$$

Now consider this DM's assessment of the lottery $q = [(100-\theta, 50), \frac{\theta}{50}; (100-\theta, 0), 1-\frac{\theta}{50}]$ generated by the choice to allocate $\theta \in [0,50]$ tokens to the other individual in task $T3$.

$$\begin{aligned} W^{EIA}(q) &= (1-\rho)\left[\frac{\theta}{50}w^{FS}(100-\theta,50) + \left(1-\frac{\theta}{50}\right)w^{FS}(100-\theta,0)\right] \\ &\quad +\rho w^{FS}(100-\theta,\theta) \\ &= (1-\rho)\left[100-\theta-\frac{\theta}{50}\beta\,(100-\theta-50) - \left(1-\frac{\theta}{50}\right)\beta(100-\theta)\right] \\ &\quad +\rho w^{FS}(100-\theta,\theta) \\ &= (1-\rho)\left[100-\theta-\beta(100-2\theta)\right] + \rho w^{FS}(100-\theta,\theta) \\ &= (1-\rho)w^{FS}(100-\theta,\theta) + \rho w^{FS}(100-\theta,\theta) = w^{FS}(100-\theta,\theta) \end{aligned}$$

Accordingly, this DM allocates the same number of tokens to the other individual in $T3$ as she does in $T1$. As the above calculations clarify, what drives this conclusion is the fact that in the EIA model ex post concerns for outcomes and ex ante concerns for opportunities serve as perfect substitutes. This is consistent with the interpretation under this model that the DM has a deep preference for fairness and reducing inequality, whether in terms of ex post outcomes or ex ante opportunities plays a symmetric role in realizing this objective.

Brock, Lange, and Ozbay (2013), however, report that when attention is restricted to non-selfish dictators, on average, more tokens were given in $T1$ than $T3$. Further, this difference is statistically significant at the 1% level. What might be the reason for this? The EM model provides one rationalization for this finding. An EM-type DM may indeed be inclined to give fewer tokens in $T3$ than in $T1$. This is because of the greater leeway that task $T3$ provides her to interpret her actions as moral compared to task $T1$. Even with a relatively fewer tokens, in task $T3$, thanks to overweighting of the probability of the socially favorable outcome $(100-\theta, 50)$, the DM can reason that her act was sufficiently moral, even if ex post the other individual ends up getting nothing. In other words, in this model, the way she assesses ex ante opportunities, specifically her expressive moral interpretation of it, may be qualitatively different from the way she assesses ex post outcomes. For instance, it can be verified numerically that with $u(x_1, x_2) = \tilde{u}(x_1) = 2\sqrt{x_1}$, $v(x_1, x_2) = \sqrt{x_2}$ and probability weighting function $\varphi(\lambda) = \exp(-b(-\ln\lambda)^a)$, where we assume that $a = 0.65$ and $b = 1.067$, the DM will give about 20 tokens in $T1$ and 15 tokens in $T3$. This shows that the EM model can accommodate behavior of DMs who are not averse to exploiting bit of moral wiggle room when it comes to interpreting their actions as moral.

Next, let us consider another set of experiments which raises the question about how much does ex post inequality matter in the presence of risk. Building on earlier findings from Bohnet et al. (2008), Bolton and Ockenfels (2010) provide experimental evidence suggesting that for decisions with social comparisons, risk taking may not depend on whether the risky option yields unequal payoffs. In their experiment, they considered a set of allocations, $A = \{(7,7), (7,16), (7,0), (9,9), (9,16), (9,0)\}$. For each allocation $(x, y) \in A$, they gave DMs in their experiment the following two choices: (i) (x, y) vs. $[(16,16), 0.5; (0,0), 0.5]$ and (ii) (x, y) vs. $[(16,0), 0.5; (0,16), 0.5]$. If ex post inequality does not matter in the presence of risk, then it should be the case that DMs' choices should not vary across (i) and (ii) in terms of whether the safe or risky alternative is chosen. This is indeed what the aggregate choice behavior in their subject pool seem to suggest. In turn, this is indicative of the fact that for many DMs, the assessment of the two allocation-lotteries $[(16,16), 0.5; (0,0), 0.5]$ and $[(16,0), 0.5; (0,16), 0.5]$ is roughly identical. Observe that this evidence can be accommodated in the EM model with $u(x_1, x_2) = \tilde{u}(x_1)$ and $v(x_1, x_2) = \tilde{v}(x_1)$. This would imply that the DM is indifferent between these two lotteries. On the other hand, a DM adhering to the EIA model always considers the lottery $[(16,16), 0.5;\ (0,0), 0.5]$ to be strictly preferable to the lottery $[(16,0), 0.5; (0,16), 0.5]$.

These set of experimental results also help clarify something important about why concerns for others' opportunities may matter for DMs with EM preferences as opposed to, say, EIA preferences. The conventional wisdom in the literature has viewed this question from the inequality aversion perspective. Specifically, the answer has been that such DMs care not just about inequality of ex post outcomes but also about inequality of ex ante opportunities as we saw in the EIA model. As such, concerns about others' opportunities enter the DM's consideration in a relative sense, i.e., in the way of making comparisons between their opportunities and her's. On the other hand, in the EM model, concerns for others' opportunities may matter too but not because of considerations about equality of opportunities. Such concerns may enter in an absolute way rather than through making relative comparisons and may capture other emotions beyond concerns for equity, including self-serving impulses like exploiting moral wiggle room.

This discussion, therefore, suggests that beyond the voting problem that is our motivation in this paper, the EM model makes certain novel contributions to the social preferences literature. At a theoretical level, it allows us to look at socially motivated behavior through the lenses of emotions distinct from those underlying the inequality aversion paradigm. What's more, this theoretical formulation can be used to validate existing experimental findings as well as suggest fresh ones.

5.4 Probability Weighting and the FGS model

In the voting model that we developed, the modeling construct of probability weighting played an important role in the analysis. Specifically, it increased the salience of voting for the moral alternative in large elections. We want to highlight now that, beyond our voting model, the concept of probability weighting can be a productive tool when it comes to studying electoral choices. To do so, we go back to the model of FGS and illustrate this point.

Recall that in the FGS setting, the population consisted of A types and B types.[18] Only B types numbering n have voting rights and, given the electoral mechanism, have a $\frac{1}{n}$ chance of influencing the outcome of the election. There are two alternatives, A and B, and a voter can vote for one of these alternatives or abstain. Choosing to vote imposes a cost of c. The *monetary* payoffs resulting for A and B types from these alternatives/choices are as follows:

	A type	B type who vote	B type who don't vote
Alternative A	$1-c$	$1-c$	1
Alternative B	0	$1+x-c$	$1+x$

[18]We present a slightly simplified version of their set-up, focusing only on the aspects that is necessary for developing the present argument.

The exact payoffs highlight why, as noted in Section 2, FGS think of A as the moral outcome in the election. Besides these monetary payoffs, FGS also allow for the possibility of two different kinds of subjective payoffs—an ethical instrumental payoff $\tilde{d}$ when the alternative A is the outcome, and an ethical expressive payoff d obtained as a consequence of voting for alternative A. We assume here that $x > \tilde{d}, d$.

We first highlight the main conclusion of the FGS model where evaluation of lotteries are done via expected utility (EU). Consider any voter and denote by q^* the probability that B is the chosen alternative when her vote is not pivotal. Then, under an EU evaluation, her payoffs from voting for A, B and not voting (denoted by $\emptyset$) are, respectively,

$$\begin{aligned}
V_{FGS}(A) &= \frac{1}{n}(1+\tilde{d}) + (1-\frac{1}{n})(q^*(1+x) + (1-q^*)(1+\tilde{d})) - c + d \\
&= \frac{1}{n}(1+\tilde{d}) + (1-\frac{1}{n})(1+\tilde{d}+q^*(x-\tilde{d})) - c + d \\
V_{FGS}(B) &= \frac{1}{n}(1+x) + (1-\frac{1}{n})(1+\tilde{d}+q^*(x-\tilde{d})) - c \\
V_{FGS}(\emptyset) &= \frac{1}{n}(1+\frac{x+\tilde{d}}{2}) + (1-\frac{1}{n})(1+\tilde{d}+q^*(x-\tilde{d}))
\end{aligned}$$

Conditional on voting, a voter votes for A over B if $V_{FGS}(A) \geq V_{FGS}(B)$. It is straightforward to verify that this is true if:

$$d \geq \frac{x-\tilde{d}}{n} \tag{1}$$

Similar calculations establish that voting for A is preferred to not voting if:

$$d - c \geq \frac{x-\tilde{d}}{2n} \tag{2}$$

Finally, a voter prefers to vote for B rather than not voting if:

$$\frac{x-\tilde{d}}{2n} \geq c$$

Accordingly, as long as $d > c$, we arrive at the conclusion that whereas a voter may prefer to vote for B over A in an election with a small electorate, in elections with a large enough electorate, she will always vote for A.

Let us now introduce probability weighting into the analysis. Perhaps the simplest way to do this is by assuming that the DM evaluates lotteries according to the rank dependent utility (RDU) model instead of using EU. Specifically, she evaluates consequences by a linear utility function and subjectively processes objective lotteries using a probability weighting function $\psi : [0,1] \to [0,1]$. Note that the 3 lotteries generated by the choice of voting for A, B and not voting are, respectively:

$$\begin{aligned}
p &= \left[1+\tilde{d}+d, \frac{1}{n}+\left(1-\frac{1}{n}\right)(1-q^*); 1+x+d, \left(1-\frac{1}{n}\right)q^*\right] \\
q &= \left[1+\tilde{d}, \left(1-\frac{1}{n}\right)(1-q^*); 1+x, \frac{1}{n}+\left(1-\frac{1}{n}\right)q^*\right] \\
r &= \left[1+\tilde{d}, \frac{1}{2n}+\left(1-\frac{1}{n}\right)(1-q^*); 1+x, \frac{1}{2n}+\left(1-\frac{1}{n}\right)q^*\right]
\end{aligned}$$

The RDU assessment of these lotteries are

$$\begin{aligned}
W(p) &= \psi\left(\left(1-\frac{1}{n}\right)q^*\right)(1+x)+\left(1-\psi\left(\left(1-\frac{1}{n}\right)q^*\right)\right)(1+\tilde{d})+d-c \\
W(q) &= \psi\left(\frac{1}{n}+\left(1-\frac{1}{n}\right)q^*\right)(1+x)+\left(1-\psi\left(\frac{1}{n}+\left(1-\frac{1}{n}\right)q^*\right)\right)(1+\tilde{d})-c \\
W(r) &= \psi\left(\frac{1}{2n}+\left(1-\frac{1}{n}\right)q^*\right)(1+x)+\left(1-\psi\left(\frac{1}{2n}+\left(1-\frac{1}{n}\right)q^*\right)\right)(1+\tilde{d})
\end{aligned}$$

Note that the RDU assessments draw on the assumption that $x > \tilde{d}$.

We can then verify that in this setting a voter prefers voting for A over B if:

$$d \geq \left[\psi\left(\frac{1}{n}+\left(1-\frac{1}{n}\right)q^*\right)-\psi\left(\left(1-\frac{1}{n}\right)q^*\right)\right](x-\tilde{d})$$

and prefers voting for A over not voting if:

$$d-c \geq \left[\psi\left(\frac{1}{2n}+\left(1-\frac{1}{n}\right)q^*\right)-\psi\left(\left(1-\frac{1}{n}\right)q^*\right)\right](x-\tilde{d})$$

The first important thing to observe is that, in contrast to the original FGS setting, a voter's optimal action depends on the actions of the other voters as q^* enters these two inequalities. Therefore, to figure out what behavior is we need to figure out the Nash equilibrium. Specifically, setting $q^* = 0$ in the above two inequalities gives us the conditions under which everyone voting for alternative A is a Nash equilibrium:

$$d \geq \psi\left(\frac{1}{n}\right)(x - \tilde{d}) \text{ and } d - c \geq \psi\left(\frac{1}{2n}\right)(x - \tilde{d})$$

Comparing these two inequalities to inequalities 1 and 2 above, shows us the difference in analysis introduced by probability weighting. It essentially tells us that the nature of subjective probability weighting in the neighborhood of zero matters for the analysis. If the probability weighting function overweights small probabilities then under the current set-up, the phenomenon of a moral bias in large elections (i.e., voting for A) requires a larger size electorate (i.e., n) than compared to the FGS benchmark. The opposite is true if there is an underweighting of probabilities around zero.

6 Comments on the Literature

Social Preferences: The first generation of social preference models (e.g., Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002)) were proposed for risk-free environments. The literature soon discovered, though, that these models cannot always be readily extended to environments of risk using standard approaches like expected utility or the available non-expected utility theories. The reason for this is that these standard models of decision making under risk all satisfy the property of stochastic dominance. On the other hand, as we have seen, social preferences under risk often violate this property. Hence, the more recent attempts in the literature have been to study social preferences while allowing for violations of stochastic dominance. Such attempts have been made by Karni and Safra (2002), Fudenberg and Levine (2012) and Saito (2013), amongst others. The decision model of social preferences presented here relates to this line of work.

Warm Glow: Our work also relates to models of warm glow. In such models, certain types of choices, typically those tied to a prosocial interpretation, provide the decision maker with a "feel good factor" that makes such choices *intrinsically* attractive to her. Because of this, warm glow payoffs may realize even when such choices do not change the social outcome; hence, they reflect a form of "impure altruism." In an influential early work, Riker and Ordeshook (1968), considered voters who derive an intrinsic warm glow payoff from the very act of voting, independent of any instrumental concerns. This allowed them to explain turnout in large elections. Andreoni, in two seminal papers, Andreoni (1989) and Andreoni (1990), uses the structure of warm glow payoffs to account for real world patterns of charitable giving and contribution to public goods. Such patterns cannot be matched in models of pure altruism. In recent years, a focus within this literature has been to develop axiomatic models of warm glow. These models serve as responses to the criticism that the earlier generation of warm glow models were by and large ad hoc. Notable contributions in this area have been made by Cherepanov, Feddersen, and Sandroni (2013), Saito (2015) and Evren and Minardi (2017). Although the broad spirit of our model of social preferences under risk is similar to classical warm glow models, ours is not a model of impure altruism. As noted earlier, expressive moral assessments in our model take the form of dependent non-instrumental valuation, i.e., such assessments are not unhinged from what end-outcomes may realize from choices.

Voting: The fact that expressive motivations in voting behavior may be a consequence of how DMs perceive the morality of their choices when these have social ramifications was first alluded to by Tullock (1971). However, to the best of our knowledge, no attempt has been made to formalize this insight within a model of social preferences that *extends beyond the immediate context of the voting problem*. Perhaps, the work that is closest to such a research agenda is the ethical voter model of Feddersen and Sandroni (2006). Their model builds on the idea of act utilitarianism and provides a theory of why, owing to ethical concerns, individuals might be motivated to vote even when the probability that their vote is pivotal is zero and there are costs associated with voting. That is, their model addresses the question of why individuals vote rather than how they vote, which is our focus here. Another paper that follows on the footsteps of this paper is Evren and Minardi (2017). It too explains turnout in large elections drawing on the presence of altruistic or ethical voters in the electorate. It should also be pointed out that the tension between instrumental and expressive motivations in voting decisions that our model highlights has also been noted by Morgan and Várdy (2012). They study a Condorcet jury model and show, in a spirit similar to our finding here, that in large elections expressive concerns may dominate instrumental ones. Unlike us, they do not focus attention on the issue of micro-foundations of expressive voting behavior.

7 Conclusion

This paper has proposed a foundation for expressive choices based on a novel theory of social preferences under risk. The decision model that we have presented formalizes the idea that, in risky social environments, decision makers who want to perceive their choices as moral or prosocial may assess the morality of their choices based on both instrumental and expressive considerations. That is, along with the concern about whether the end-outcome that their choice facilitates is socially favorable or not, such decision makers may also be interested in reasoning whether their choice per se, distinct from the outcome that results from it, can be interpreted as moral.

The key analytical ingredient underlying our decision model is the difference in how instrumental and expressive moral assessments are made. Expressive moral assessments are characterized by a departure from the Bayesian calculus. We modelled this by using probability weighting functions which allowed decision makers to subjectively weight objective probabilities of events. Especially, we saw that the tendency to over-weight small probabilities and under-weight large ones (of morally favorable outcomes occurring)—a phenomenon known as regressive probability weighting—played a crucial part in the analysis.

We have used our decision model to highlight how, when applied to an electoral setting, an expressive consideration in voting behavior endogenously emerges. We have shown how voting behavior that may be interpreted as being against one's self-interest can be rationally understood as emerging out of an interaction between instrumental and expressive concerns that our decision model highlights. Specifically, owing to regressive probability weighting, we have shown how in large elections expressive concerns may dominate instrumental ones. This, in turn, helped us shed further light on why such elections may exhibit a moral bias.

Appendix

Proof (Proposition 4.2)

As in the text, let λ denote the probability that the RV's vote is pivotal. Under the mechanism that determines the electoral outcome, $\lambda = \frac{1}{n}$, where n is the size of the electorate. Further, let γ denote the probability that alternative A is the electoral outcome when her vote is not pivotal. Of course, γ is an endogenous object and determined by the strategies of all other voters. Then, the probability distributions over electoral outcomes generated by the RV voting for alternatives A and B, respectively, are:

$$\begin{aligned} p^1 &= [A, \lambda + (1-\lambda)\gamma; B, 1 - \lambda - (1-\lambda)\gamma] \\ p^2 &= [A, (1-\lambda)\gamma; B, 1 - (1-\lambda)\gamma] \end{aligned}$$

Under an EM representation of the RV's preferences, these two lotteries are evaluated as:[19]

$$\begin{aligned} W(p^1) &= u_H + v_L + (\lambda + (1-\lambda)\gamma)(u_L - u_H) \\ &\quad + [\sigma\varphi(\lambda + (1-\lambda)\gamma) + (1-\sigma)(\lambda + (1-\lambda)\gamma)](v_H - v_L) \\ W(p^2) &= u_H + v_L + (1-\lambda)\gamma(u_L - u_H) \\ &\quad + [\sigma\varphi((1-\lambda)\gamma) + (1-\sigma)(1-\lambda)\gamma](v_H - v_L) \end{aligned}$$

Accordingly,

$$\begin{aligned} &W(p^2) - W(p^1) = \\ &\quad (v_H - v_L)\left[\lambda\left(\frac{u_H - u_L}{v_H - v_L} - (1-\sigma)\right) - \sigma(\varphi(\lambda + (1-\lambda)\gamma) - \varphi((1-\lambda)\gamma))\right] \end{aligned}$$

Letting $\nu = \frac{u_H - u_L}{v_H - v_L}$, it follows that voting for alternative B is a best response for the RV iff:

$$W(p^2) - W(p^1) \geq 0 \Leftrightarrow g(\lambda, \gamma) := \lambda(\nu - 1 + \sigma) - \sigma[\varphi(\lambda + (1-\lambda)\gamma) - \varphi((1-\lambda)\gamma)] \geq 0,$$

[19] These expressions follow after simplifying the EM assessments of the lotteries p^1 and p^2 that are given below:
$W(p^1) = (\lambda + (1-\lambda)\gamma)[u_L + v_H + \sigma(\varphi(\lambda + (1-\lambda)\gamma)v_H + (1 - \varphi(\lambda + (1-\lambda)\gamma))v_L - v_H)] + (1 - (\lambda + (1-\lambda)\gamma))[u_H + v_L + \sigma(\varphi(\lambda + (1-\lambda)\gamma)v_H + (1 - \varphi(\lambda + (1-\lambda)\gamma))v_L - v_L)]$
$W(p^2) = (1-\lambda)\gamma[u_L + v_H + \sigma(\varphi((1-\lambda)\gamma)v_H + (1 - \varphi((1-\lambda)\gamma))v_L - v_H)] + (1 - (1-\lambda)\gamma)[u_H + v_L + \sigma(\varphi((1-\lambda)\gamma)v_H + (1 - \varphi((1-\lambda)\gamma))v_L - v_L)]$

where $g : [0,1]^2 \to \mathbb{R}$. It is worth noting that the function g is continuous, since the function φ is continuous.

We now establish the first part of the proposition, i.e., when the electorate is above a particular size, everyone voting for alternative A is the unique symmetric pure strategy Nash equilibrium. Suppose everyone other than the RV votes for alternative A, i.e., $\gamma = 1$. We show below that, in this case, the RV's best response is to vote for A as well. First, consider the g function and observe that,

$$g(\lambda, 1) = \lambda(\nu - 1 + \sigma) - \sigma(1 - \varphi(1 - \lambda)).$$

Further, when $\lambda = 0$, $g(0, 1) = 0$. Next, note that

$$\frac{dg(\lambda, 1)}{d\lambda} = \nu - 1 + \sigma - \sigma\varphi'(1 - \lambda) = \sigma\left[\frac{\nu - 1}{\sigma} + 1 - \varphi'(1 - \lambda)\right].$$

So, by assumption $[A1]$, for $\lambda \in [0, \hat{\mu}]$, $\frac{dg(\lambda,1)}{d\lambda} < 0$, i.e., $g(., 1)$ is strictly decreasing in λ in the interval $[0, \hat{\mu}]$. It, therefore, follows that for any $\lambda \in (0, \hat{\mu}]$, $g(\lambda, 1) < 0$, i.e., the RV's best response is to vote for alternative A. Now, let $\overline{n}$ be the smallest integer such that $\overline{n} \geq \frac{1}{\hat{\mu}}$. Then, for all $n \geq \overline{n}$, everyone voting for alternative A is a Nash equilibrium. Now, consider the case when everyone other than the RV votes for alternative B, i.e., $\gamma = 0$. Note that, in this case,

$$g(\lambda, 0) = \lambda(\nu - 1 + \sigma) - \sigma(\varphi(\lambda) - \varphi(0)) = \lambda(\nu - 1 + \sigma) - \sigma\varphi(\lambda).$$

Further, verify that,

$$\frac{dg(\lambda, 0)}{\partial\lambda} = \sigma\left[\frac{\nu - 1}{\sigma} + 1 - \varphi'(\lambda)\right].$$

So, again by $[A1]$, for $\lambda \in [0, \hat{\mu}]$, $\frac{dg(\lambda,0)}{d\lambda} < 0$. Since $g(0, 0) = 0$, it follows that for any $\lambda \in (0, \hat{\mu}]$, $g(\lambda, 0) < 0$, i.e., the RV's best response is to vote for alternative A. Accordingly, for $n \geq \overline{n}$, everyone voting for alternative B is not a Nash equilibrium. Hence, whenever $n \geq \overline{n}$, everyone voting for alternative A is the unique symmetric pure strategy Nash equilibrium.

We next establish the second part of the proposition. Suppose everyone in the electorate other than the RV is voting for alternative B, i.e., $\gamma = 0$. We show below that, in this case, the RV's best response is to vote for B as well. As seen above, $g(\lambda, 0) = \lambda(\nu - 1 + \sigma) - \sigma\varphi(\lambda)$. Since $\nu > 1$ and $\varphi(1) = 1$, it follows that $g(1, 0)$

$= \nu - 1 > 0$. By continuity of g, it follows that there exists an interval $(\lambda_1, 1]$, such that for all $\lambda \in (\lambda_1, 1]$, $g(\lambda, 0) > 0$, i.e., the RV's best response is to vote for alternative B. Accordingly, everyone voting for alternative B is a symmetric Nash equilibrium when $\lambda \in (\lambda_1, 1]$. Next, consider the case when everyone other than the RV is voting for alternative A, i.e., $\gamma = 1$. In this case, as seen above, $g(\lambda, 1) = \lambda(\nu - 1 + \sigma) - \sigma(1 - \varphi(1 - \lambda))$ and, in particular, $g(1, 1) = \nu - 1 > 0$. Once again, by the continuity of g, it follows that there exists an interval $(\lambda_2, 1]$, such that for all $\lambda \in (\lambda_2, 1]$, $g(\lambda, 1) > 0$, i.e., the RV's best response is to vote for alternative B. Accordingly, everyone voting for alternative A is not a Nash equilibrium when $\lambda \in (\lambda_2, 1]$. Let $\lambda'' = \max\{\lambda_1, \lambda_2\}$, and let $\underline{n}$ be the largest positive integer such that $\underline{n} \leq \frac{1}{\lambda''}$. It follows that for all $n \leq \underline{n}$, everyone voting for alternative B is the unique symmetric pure strategy Nash equilibrium.

Proof (Proposition 5.1)

Consider a society of size n and the voting choice problem of a given individual i in this society. Suppose all other individuals in this society are adopting a symmetric cutoff strategy of the type stated in Proposition 5.1, i.e., all individuals below the cutoff cost level vote for A and all others abstain. We show below that for societies that are large enough (i.e., n greater than some $\bar{n}$), the best response of this individual is to adopt a similar cutoff strategy. In the process, we derive what the cutoff c_n^* has to be in such a Bayes-Nash equilibrium.

To that end, suppose that all individuals other than i whose costs are below some level $c \in [0, \bar{c}]$ vote for A; and all others abstain. Then the lotteries over electoral outcomes generated by i's choice of voting for A, voting for B and abstaining are, respectively,[20]

$$q^1 = \delta_A; q^2 = [A, 1 - \pi_n(c); B, \pi_n(c)]; q^3 = [A, 1 - \pi'_n(c); B, \pi'_n(c)]$$

where,

$$\begin{aligned} \pi_n(c) &= (1 - F(c))^{n-1} + (n-1)F(c)(1 - F(c))^{n-2} \\ \pi'_n(c) &= (1 - F(c))^{n-1} \end{aligned}$$

[20] If i votes for A, then A wins for sure—δ_A refers to the degenerate lottery that results in A with probability 1. If she votes for B, then B wins if no more than one of the other $n-1$ individuals have cost less that c, otherwise A wins. If she abstains, then B wins if none of the other $n-1$ individuals have cost less that c, otherwise A wins.

We first highlight some properties of the functions $\pi_n : [0, \bar{c}] \to [0, 1]$ and $\pi'_n : [0, \bar{c}] \to [0, 1]$. Observe that, for any n, both these functions are continuous, since $F(c)$ is differentiable and, hence, continuous. Further, they converge pointwise to the constant function $\tilde{\pi}(c) = 0$ for all $c \in [0, \bar{c}]$. That is, $\lim_{n\to\infty}(\pi_n(c), \pi'_n(c)) = (0, 0)$, for all $c \in [0, \bar{c}]$.[21] What we now establish is that this convergence is, in fact, uniform.

Lemma. $\pi_n : [0, \bar{c}] \to [0, 1]$ and $\pi'_n : [0, \bar{c}] \to [0, 1]$ converge uniformly to $\tilde{\pi} : [0, \bar{c}] \to [0, 1]$.

Proof. The uniform convergence of π'_n to $\tilde{\pi}$ follows immediately from Dini's Theorem, since (π'_n) is a monotone sequence of continuous functions on $[0, \bar{c}]$ that converges to $\tilde{\pi} : [0, \bar{c}] \to [0, 1]$. To establish the uniform convergence of π_n, it suffices to show that the function $\pi''_n(c) = (n - 1)F(c)(1 - F(c))^{n-2}$, $c \in [0, \bar{c}]$, converges uniformly to $\tilde{\pi}$. To do so, first, note that for any $c \in [0, \bar{c}]$, there exists $\hat{n}(c) = \left\lceil \frac{1}{F(c)} \right\rceil \in \mathbb{N}$ such that the tail sequence $(\pi''_n(c))_{n>\hat{n}(c)}$ is a decreasing one.[22] This implies that, letting $\hat{n} = \hat{n}(0)$, the tail sequence $(\pi''_n)_{n>\hat{n}}$ is a decreasing one. Next, consider any $\epsilon > 0$. Observe that for any $c \in [0, \bar{c}]$, there exists $n_c > \hat{n}$ s.t., $0 \le \pi''_{n_c}(c) < \frac{\epsilon}{2}$. Further, since π''_{n_c} is continuous, there exists $\chi_c > 0$, s.t., for all $c' \in I(c, \chi_c) := \{c'' \in [0, \bar{c}] : |c - c''| < \chi_c\}$, $|\pi''_{n_c}(c') - \pi''_{n_c}(c)| < \frac{\epsilon}{2}$. It therefore follows that for all $c' \in I(c, \chi_c)$,

$$0 \le \pi''_{n_c}(c') \le |\pi''_{n_c}(c') - \pi''_{n_c}(c)| + \pi''_{n_c}(c) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Further, since $(\pi''_n)_{n\ge\hat{n}}$ is a decreasing sequence, $0 \le \pi''_n(c') < \epsilon$, for all $n > n_c$. Finally, note that $(I(c, \chi_c))_{c\in[0,\bar{c}]}$ is an open cover of $[0, \bar{c}]$, which, being compact, implies that there exists a finite subcover $\{I(c_1, \chi_{c_1}), \ldots, I(c_K, \chi_{c_K})\}$. Let $\underline{n} = \max\{n_{c_1}, \ldots, n_{c_K}\}$. Then, for all $n \ge \underline{n}$, and all $c \in [0, \bar{c}]$, $0 \le \pi''_n(c) < \epsilon$ □

Next, note that for any lottery $q = [A, 1 - \pi; B, \pi]$, the EM assessment of the lottery simplifies to:

$$W(q; \pi) = u_L + v_L + \pi(u_H - u_L) + [\sigma\varphi(1 - \pi) + (1 - \sigma)(1 - \pi)](v_H - v_L)$$

This implies that:[23]

$$\frac{\partial W}{\partial \pi} = \sigma(v_H - v_L)\left[1 + \frac{\nu - 1}{\sigma} - \varphi'(1 - \pi)\right]$$

[21]The fact that $\lim_{n\to\infty}(\pi_n(0), \pi'_n(0)) = (0, 0)$ draws on the assumption that $F(0) > 0$.

[22]For any real number r, $\lceil r \rceil$ refers to the smallest integer greater than or equal to r.

[23]Recall that $\nu = \frac{u_H - u_L}{v_H - v_L}$.

So, by assumption $[A1]$, it follows that for $\pi < \hat{\mu}$, we have $\varphi'(1-\pi) > 1 + \frac{\nu-1}{\sigma}$ and, accordingly, $\frac{\partial W}{\partial \pi} < 0$. That is increasing π in that range lowers the EM assessment of such a lottery.

Next, note that by the uniform convergence of the sequences (π_n) and (π'_n) established in the lemma above, we can find $\bar{n}$ such that for all $n \geq \bar{n}$, $\pi'_n(c) < \pi_n(c) < \hat{\mu}$, for all $c \in [0, \bar{c}]$. Accordingly, $W(q^3) > W(q^2)$, for all $n \geq \bar{n}$ and $c \in [0, \bar{c}]$. That is, for $n \geq \bar{n}$, the relevant comparison for this voter is between voting for A and abstaining, no matter what the cutoff c adopted by the other voters is. So, let $h(\pi'_n(c)) = W(q^1) - W(q^3)$. Elementary calculations imply that:

$$h(\pi'_n(c)) = (v_H - v_L)[1 - \sigma\varphi(1 - \pi'_n(c)) - (1 - \sigma)(1 - \pi'_n(c)) - \pi'_n(c)\nu]$$

Accordingly,

$$\frac{dh}{d\pi'_n} = \sigma(v_H - v_L)\left[\varphi'(1 - \pi'_n(c)) - \left(1 + \frac{\nu - 1}{\sigma}\right)\right]$$

For $n \geq \bar{n}$, $\pi'_n(c) < \hat{\mu}$, for all $c \in [0, \bar{c}]$. Hence, $\varphi'(1-\pi'_n(c)) > 1 + \frac{\nu-1}{\sigma}$ and $\frac{dh}{d\pi'_n} > 0$. Given that h is continuous and $h(0) = 0$, it follows that for $\pi'_n(c) < \hat{\mu}$, $h(\pi'_n(c)) > 0$, i.e., $W(q^1) > W(q^3)$, for all $n \geq \bar{n}$ and $c \in [0, \bar{c}]$.

Now, for any $n \geq \bar{n}$, define the function $H_n : [0, \bar{c}] \to \mathbb{R}$ by $H_n(c) = h(\pi'_n(c))$. A symmetric Bayes Nash equilibrium of the type stated in the Proposition exists for any such n if there exists $c^*_n \in [0, \bar{c}]$ such that $H_n(c^*_n) = c^*_n$. Note that $H_n(0) > 0$, $\frac{dH_n(c)}{dc} = \frac{dh}{d\pi'_n}\frac{d\pi'_n(c)}{dc} < 0$ and $H_n(\bar{c}) = 0$.[24] Therefore, there exists a unique $c^*_n \in (0, \bar{c})$ satisfying $H_n(c^*_n) = c^*_n$, establishing that all voters following the strategy

$$s(c) = \begin{cases} \textit{vote } A, & \text{if } c \leq c^*_n \\ \textit{abstain}, & \text{if } c > c^*_n \end{cases}$$

is a Bayes-Nash equilibrium for this n person society.

Proof (Proposition 5.2)

For any voter, $w^{FS}(x_B, y_B) > w^{FS}(x_A, y_A)$ implies that

$$\begin{aligned} x_B - \beta'(x_B - y_B) &> x_A - \beta'(x_A - y_A), \text{ where } \beta' := n'\beta \\ \Rightarrow (1 - \beta')(x_B - x_A) &> \beta'(y_A - y_B) \end{aligned}$$

[24]Note that $\frac{d\pi'_n(c)}{dc} = -(n-1)(1 - F(c))^{n-2}F'(c) < 0$

Next, like in the proof of Proposition 4.2, consider the lotteries p^1 and p^2 generated by a representative voter voting for alternatives A and B, respectively:

$$\begin{aligned} p^1 &= [(x_A, y_A), \lambda + (1-\lambda)\gamma; (x_B, y_B), 1 - \lambda - (1-\lambda)\gamma] \\ p^2 &= [(x_A, y_A), (1-\lambda)\gamma; (x_B, y_B), 1 - (1-\lambda)\gamma] \end{aligned}$$

Here, $\lambda = \frac{1}{n}$ is the probability that this voter's vote is pivotal and γ denotes the probability that alternative A is the electoral outcome when her vote is not pivotal. Note that expected payoffs to voters and non-voters under p^1 and p^2 are, respectively,

$$\begin{aligned} E_{p^1}(x) &= (x_B - (\lambda + (1-\lambda)\gamma)(x_B - x_A), y_B + (\lambda + (1-\lambda)\gamma)(y_A - y_B)) \\ E_{p^2}(x) &= (x_B - (1-\lambda)\gamma(x_B - x_A), y_B + (1-\lambda)\gamma(y_A - y_B)) \end{aligned}$$

After a bit of simplification, it can be shown that any such voter's EIA assessment of the two lotteries are given by:[25]

$$\begin{aligned} W^{EIA}(p^1) &= x_B - \beta'(x_B - y_B) - (\lambda + (1-\lambda)\gamma)(x_B - x_A) \\ &\quad + \beta'(\lambda + (1-\lambda)\gamma)(x_B - x_A + y_A - y_B) \\ W^{EIA}(p^2) &= x_B - \beta'(x_B - y_B) - (1-\lambda)\gamma(x_B - x_A) \\ &\quad + \beta'(1-\lambda)\gamma(x_B - x_A + y_A - y_B) \end{aligned}$$

Accordingly,

$$\begin{aligned} W^{EIA}(p^2) - W^{EIA}(p^1) &= \lambda(x_B - x_A) - \beta'\lambda(x_B - x_A + y_A - y_B) \\ &= \lambda[(1-\beta')(x_B - x_A) - \beta'(y_A - y_B)] > 0 \end{aligned}$$

That is, irrespective of what λ and, hence, n is, such a voter always votes for alternative B.

[25] $W^{EIA}(p^1) = (1-\rho)[(\lambda + (1-\lambda)\gamma)(x_A - \beta'(x_A - y_A)) + (1 - (\lambda + (1-\lambda)\gamma))(x_B - \beta'(x_B - y_B))] + \rho[x_B - (\lambda + (1-\lambda)\gamma)(x_B - x_A) - \beta'(x_B - y_B - (\lambda + (1-\lambda)\gamma)(x_B - x_A + y_A - y_B)] = (1-\rho)[x_B - (\lambda + (1-\lambda)\gamma)(x_B - x_A) - \beta'(\lambda + (1-\lambda)\gamma)(x_A - y_A - x_B + y_B) - \beta'(x_B - y_B)] + \rho[x_B - (\lambda + (1-\lambda)\gamma)(x_B - x_A) - \beta'(x_B - y_B) + \beta'(\lambda + (1-\lambda)\gamma)(x_B - x_A + y_A - y_B)] = x_B - \beta'(x_B - y_B) - (\lambda + (1-\lambda)\gamma)(x_B - x_A) + \beta'(\lambda + (1-\lambda)\gamma)(x_B - x_A + y_A - y_B)$
$W^{EIA}(p^2) = (1-\rho)[(1-\lambda)\gamma(x_A - \beta'(x_A - y_A)) + (1 - (1-\lambda)\gamma)(x_B - \beta'(x_B - y_B))] + \rho[x_B - (1-\lambda)\gamma)(x_B - x_A) - \beta'(x_B - y_B - (1-\lambda)\gamma(x_B - x_A + y_A - y_B)] = (1-\rho)[x_B - (1-\lambda)\gamma(x_B - x_A) - \beta'(1-\lambda)\gamma(x_A - y_A - x_B + y_B) - \beta'(x_B - y_B)] + \rho[x_B - (1-\lambda)\gamma(x_B - x_A) - \beta'(x_B - y_B) + \beta'(1-\lambda)\gamma(x_B - x_A + y_A - y_B)] = x_B - \beta'(x_B - y_B) - (1-\lambda)\gamma(x_B - x_A) + \beta'(1-\lambda)\gamma(x_B - x_A + y_A - y_B)$

References

Andreoni, James. 1989. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of political Economy* 97 (6):1447–1458.

———. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The economic journal* 100 (401):464–477.

Bohnet, Iris, Fiona Greig, Benedikt Herrmann, and Richard Zeckhauser. 2008. "Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States." *American Economic Review* 98 (1):294–310.

Bolton, Gary E. and Alex Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90:166–193.

Bolton, Gary E and Axel Ockenfels. 2010. "Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States: Comment." *American Economic Review* 100 (1):628–33.

Brennan, Geoffrey and James Buchanan. 1984. "Voter Choice: Evaluating Political Alternatives." *American Behavioral Scientist* 28:185–201.

Brennan, Geoffrey and Loren Lomasky. 1984. "Inefficient Unanimity." *Journal of Applied Philosophy* 1 (1):151–163.

———. 1993. *Democracy and Decision*. Cambridge: Cambridge University Press.

Brock, J. Michelle, Andreas Lange, and Erkut Y. Ozbay. 2013. "Dictating the Risk: Experimental Evidence on Giving in Risky Environments." *American Economic Review* 103:415–437.

Buchanan, James. 1954. "Individual Choice in Voting and the Market." *Journal of Political Economy* 62:334–343.

Charness, Gary and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics* 117:817–869.

Cherepanov, Vadim, Tim Feddersen, and Alvaro Sandroni. 2013. "Revealed preferences and aspirations in warm glow theory." *Economic Theory* 54 (3):501–535.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.

Evren, Özgür and Stefania Minardi. 2017. "Warm-glow Giving and Freedom to be Selfish." *The Economic Journal* 127 (603):1381–1409.

Feddersen, Timothy, Sean Gailmard, and Alvaro Sandroni. 2009. "Moral Bias in Large Elections: Theory and Experimental Evidence." *American Political Science Review* 103:175–192.

Feddersen, Timothy and Alvaro Sandroni. 2006. "A Theory of Participation in Elections." *The American Economic Review* 96:1271–1282.

Fehr, Ernst and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics* 114:817–868.

Fiorina, Morris P. 1976. "The Voting Decision: Instrumental and Expressive Aspects." *The Journal of Politics* 38:390–413.

Fudenberg, Drew and David K. Levine. 2012. "Fairness, Risk Preferences and Independence: Impossibility Theorems." *Journal of Economic Behavior and Organization* 81:602–612.

Ghirardato, Paolo and Massimo Marinacci. 2002. "Ambiguity Made Precise: A Comparative Foundation." *Journal of Economic Theory* 102:251–289.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47:263–291.

Karni, Edi and Zvi Safra. 2002. "Individual Sense of Justice: A Utility Representation." *Econometrica* 70:263–284.

Krawczyk, Michal and Fabrice LeLec. 2010. "Give Me a Chance! An Experiment in Social Decision under Risk." *Experimental Economics* 13:500–511.

Morgan, John and Felix Várdy. 2012. "Mixed Motives and the Optimal Size of Voting Bodies." *Journal of Political Economy* 120:986–1026.

Prelec, Drazen. 1998. "The Probability Weighting Function." *Econometrica* 66:497–527.

Riker, William H and Peter C Ordeshook. 1968. "A Theory of the Calculus of Voting." *The American Political Science Review* 62:25–42.

Saito, Kota. 2013. "Social Preferences Under Risk: Equality of Opportunity versus Equality of Outcome." *American Economic Review* 103:3084–3101.

———. 2015. "Impure altruism and impure selfishness." *Journal of Economic Theory* 158:336–370.

Sandbu, Martin E. 2007. "Valuing Processes." *Economics and Philosophy* 23:205–235.

Schuessler, Alexander A. 2000. *A Logic of Expressive Choice*. Princeton: Princeton University Press.

Tullock, Gordon. 1971. "The Charity of the Uncharitable." *Western Economic Journal* 9:379–392.

Wakker, Peter P. 2010. *Prospect Theory for Risk and Ambiguity*. Cambridge: Cambridge University Press.